NASA CR- 166706

(NASA-CR-166706) LAMMR WORLD DATA BASE DOCUMENTATION SUPPORT AND DEMONSTRATIONS Final Report (Business and Technological Systems, Inc.) 42 p HC A03/MF A01 CSCL 05B

N81-32073

Unclas
G3/82 35024

BUSINESS AND TECHNOLOGICAL SYSTEMS, INC.

BTS

BTS-FR-80-129

October, 1980

FINAL REPORT FOR LAMMR WORLD DATA BASE DOCUMENTATION SUPPORT AND DEMONSTRATIONS

by

Roland Chin

Paul Beaudet

submitted to

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Goddard Space Flight Center

Greenbelt, Maryland 20771

in response to

RFP No. 5-65568/337

Contract No. NAS 5-26239

by

BUSINESS AND TECHNOLOGICAL SYSTEMS, INC.

Aerospace Building, Suite 440

10210 Greenbelt Road

Seabrook, Maryland 20801

TABLE OF CONTENTS

LIST OF FIGURES

1.0 INTRODUCTION AND SUMMARY

The primary purpose of the World Surface Map is to provide the LAMMR subsystem with world surface type classifications that are used to set up LAMMR Level II process control. This data base will be accessed solely by the LAMMR subsystem. The SCATT and ALT subsystems will access the data base indirectly through the T_b (Brightness Temperature) Data Bank, where the surface types have been updated from a priori to current classification, and where the surface types have already been organized on an orbital subtrack basis.

The single most important factor in the design of the World Surface Maps is the ease of access to the information while the complexity of generating these maps is of lesser importance because their generation is a one-time, off-line process. The World Surface Map provides storage of information with a resolution of 7 km necessary to set flags concerning the earth's features with a different set of maps for each month of the year.

- Land;
- First year ice;
- Multi-year ice;
- RFI at 6.6 GHz, 5.1 GHz and 4.3 GHz, and
- Ocean.

Properties of the World Surface data base of importance are:

- a "telescoping" scheme to minimize total data storage,
- geographic contiguity,
- access efficiency, and
- uniform distribution of cell centers.

Several earth-oriented data base structures have been conceptualized, in numerous studies (References 1, 2, and 3), including the spherical icosahedron and the spherical cube. These various methods are considered for adaptation to the LAMMR World Surface Map data base structure. Indexing of these data bases, transformation between geographic coordinates and computer storage indexing, data accessing, and retrieval have been examined.

Results of this study show that a latitude-longitude indexing system is most suitable for the World Surface Map. This indexing system eliminates both the necessity to project a geometrical structure onto a circumscribing sphere, and the need to explicitly compute indices for every cell. Thus, retrieval is simple and rapid. The LAMMR data is in instrument coordinates, with latitude-longitude addressing tags accompanying the data.

The main features of the World Surface Maps are summarized as follows:

- The earth's surface is divided into a number of geographically-contiguous regions, each associated with a data base record.

- The addressing scheme is binary sequential. This permits "telescoping" whereby regions of diminishing size are addressed by adding bits to indices of increasing length.

- The cell addresses are binary coded latitude-longitude. Conversion from one coordinate system to another is not necessary.

To achieve approximately 7 km gridding intervals a total storage of about 16 million cells is required for each of the latitude-longitude based World Surface Maps. The earth's surface is partitioned into $2N^2$ squares with N equidistant partitions on the latitude axis and $2N$ equidistant partitions on the longitude axis. Each square is considered a logical record consisting of $M \times M$ cells. Each cell has a gridding interval of 7 km along the latitude axis. The gridding interval along the Equator is also 7 km, but for any other parallel of latitude, the gridding interval is less than 7 km. Results of the study reported in Section 2 show that $N = 256$ and $M = 12$ is a good choice and N can be further broken down into telescoping levels to reduce the amount of storage required.

Seventeen bits are required for indexing one of the $2 \times 256^2 = 131072$ regions. Another 8 bits are used for addressing cells within the region. Coordinate transformations are not necessary because both the geographic coordinates and the data base indices are the same. However, there is a requirement to encode latitude-longitude coordinates into binary addresses.

By incorporating telescoping into the data base retrieving process, the amount of storage required is reduced. A rough estimate of 25% reduction in storage is possible; the data base is thereby reduced to 12 million units per map. Four bits per flag are needed for storage of the information. The total storage for 12 such maps is 576 million bits.

It should be noted here that a geographically-referenced World Brightness Map may be required for data comparisons between different spacecrafts or instruments as well as data gathering for research purposes. The World Brightness Map requires cells of equal area and minimal distortion. This requirement is most likely to be fulfilled by spherical icosahedron structure, the details which are discussed in Section 3.

1.1 Future Study and Recommendations

The World Surface Map needed to provide the surface type classification flags has been organized in area blocks within area blocks in a telescopic fashion. The purpose of this is to provide an efficient means of access to this data bank and to reduce the total data volume required. Total data volume requirements and efficient access schemes are important considerations in the design of the LAMMR subsystem algorithms. Additional study is required to determine whether the selected method or some other such method is most applicable. At the same time, with the advancing computational and storage capabilities of ground systems, such as will be implemented for the PPF (Primary Processing Facility), data volume might play a less significant role than anticipated. As a result, the telescopic approach to the World Surface Map may turn out to be not worth the overhead associated with it. This question can only be firmly resolved after the hardware system of the PPF has been defined.

2.0 DATA BASE BASED ON LATITUDE-LONGITUDE COORDINATES

This section presents results of the latitude-longitude coordinates Earth data base study. The gridding system of this data base is based on the latitude-longitude geographical coordinate reference system on the spherical Earth. In this coordinate, the latitude and longitude are treated as the X and Y axis respectively by mapping the Earth on a plane. The Earth as a plane reference surface with latitude-longitude gridding system is shown in Figure 2-1. The gridding intervals along the X-axis, the meridian, are equal. At first sight this may seem to be true along the Y-axis too, but the actual gridding intervals along the Y-axis close to the poles are smaller than those close to the Equator. It is obvious that this data base structure does not have equal-area and evenly distributed elements.

The simplicity of the coordinate system allows such a data base to be capable of quick and direct-retrieval of data with minimal computations. All data in the data base may be indexed by directly calculated addresses. The determination of data base addresses is straightforward. Once the latitude-longitude coordinates are available, the transformation to the rectangular storage coordinates is just a matter of relabelling the coordinates. The binary serial address can then be obtained by simple table look-up procedures.

2.1 Telescoping Addressing Scheme

The specific addressing scheme discussed in this section is based on the binary division scheme. The Earth's surface is divided into a number of square regions. Each region is further divided, to the requisite resolution level, by a two-dimensional binary grid, as shown in Figure 2-2. On each level of division, the areas are divided into quadrants, which are labeled by a 2-bit binary number. Each level of division is indicated by the addition of two binary bits to the least-significant end of the address. This binary address is the serial location of a point in

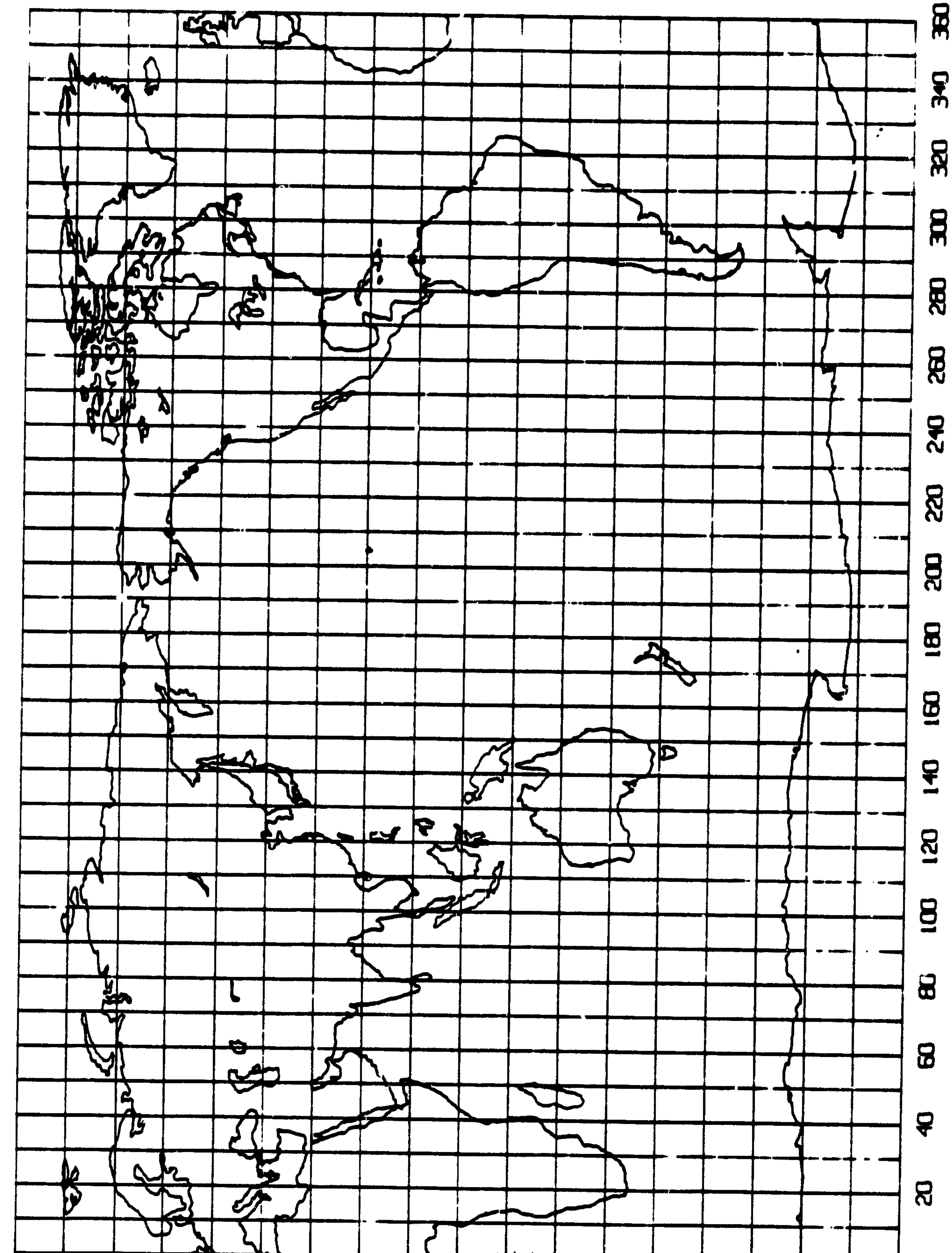


Figure 2-1
World Map with Latitude-Longitude Coordinate System

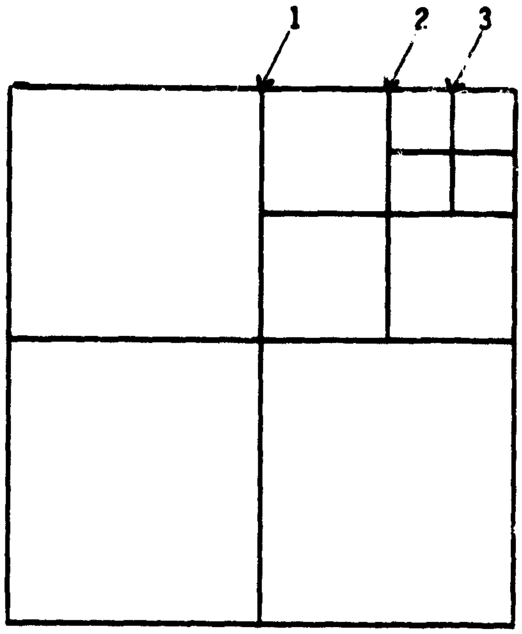


Figure 2-2 Binary Division Scheme

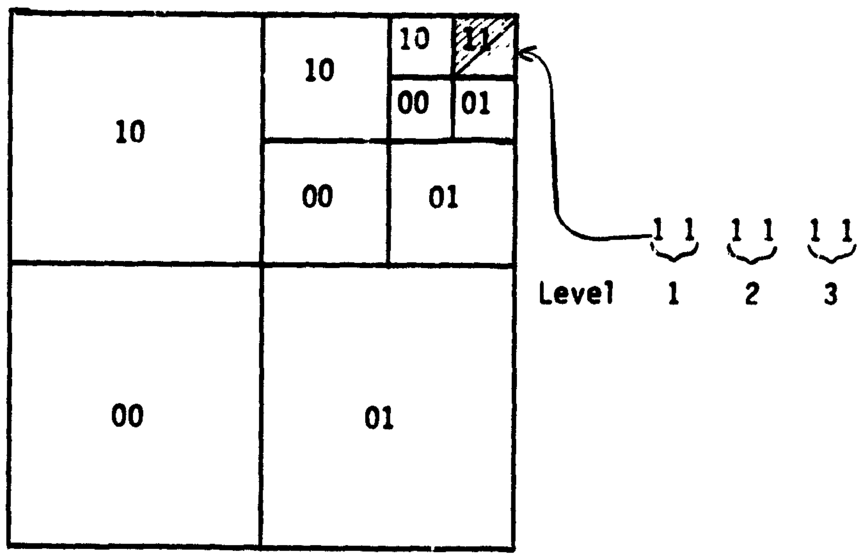


Figure 2-3 Illustrative Addressing Within a Region by Binary Bits

the square region. An address to the third-level of division is presented in Figure 2-3 as an example. The expression of an address as a single bit string allows storage of addresses as single machine words and the expandibility and generality of the serial string permits the use at any resolution level by "telescoping".

The surface flags are organized in area blocks within area blocks in a telescopic fashion. The main purpose is to reduce the total data volume required. This is done by examining the homogeneity of large area blocks. For example, a large area block over land only requires the storage of a single surface flag. In one experiment, the number of homogeneous area blocks over land was determined for different block sizes. A 25% reduction in data volume was shown to be possible. Figure 2-4 shows the plot of the heuristic data.

Telescoping procedure can be thought of as a tree-like structure. The procedure involves a sequence of partitioning area blocks into finer resolution. And, it is characterized by the fact that each partitioning step is subjected to examinations and decisions before further partitioning continues. When the structure is diagrammed to show the hierarchy among the procedure, it exhibits a characteristic tree-like aspect.

Figure 2-5 is an example to show the conceptual telescoping process. In this example, the address of a resolution element is broken down into four parts. The first part of the address, A_1 , points to a large area block labelled AREA1. The bit string A_2 addresses a smaller area block AREA2 in which AREA2 is contained in AREA1. Similarly, A_3 and A_4 address AREA3 and CELL, respectively. In actual storage, AREA3 might be structured as a logical record containing resolution cells. At each of the telescoping steps, a decision is made based on the homogeneity of the area examined. The process continues if the area block is heterogeneous. This corresponds to the continuous processing of the tree at a lower level. The addressing process ends if the area block being examined is homogeneous or the tree level with the desired resolution is reached.

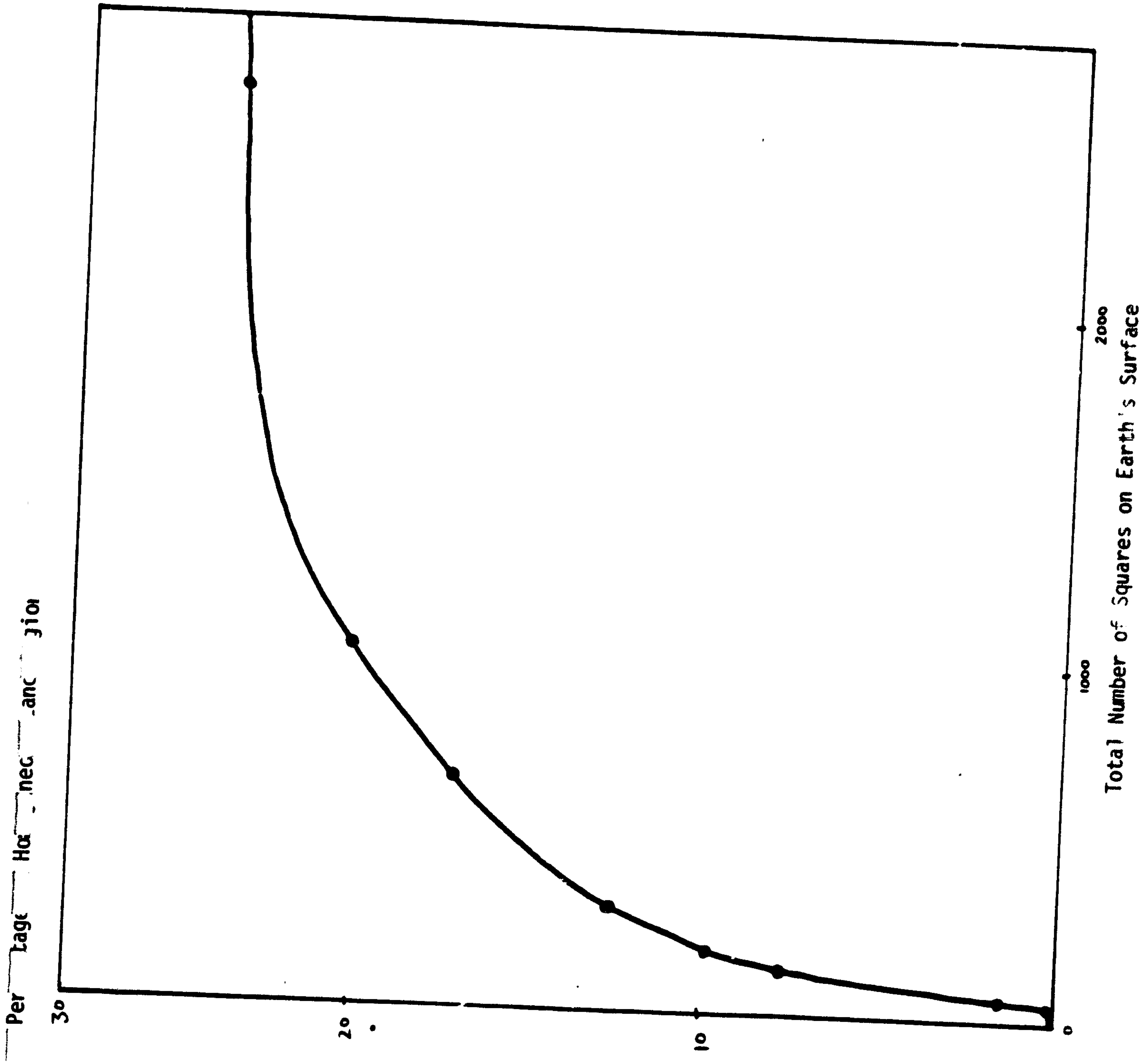


Figure 2-4 Percentage of Homogeneous Area Blocks Over Land

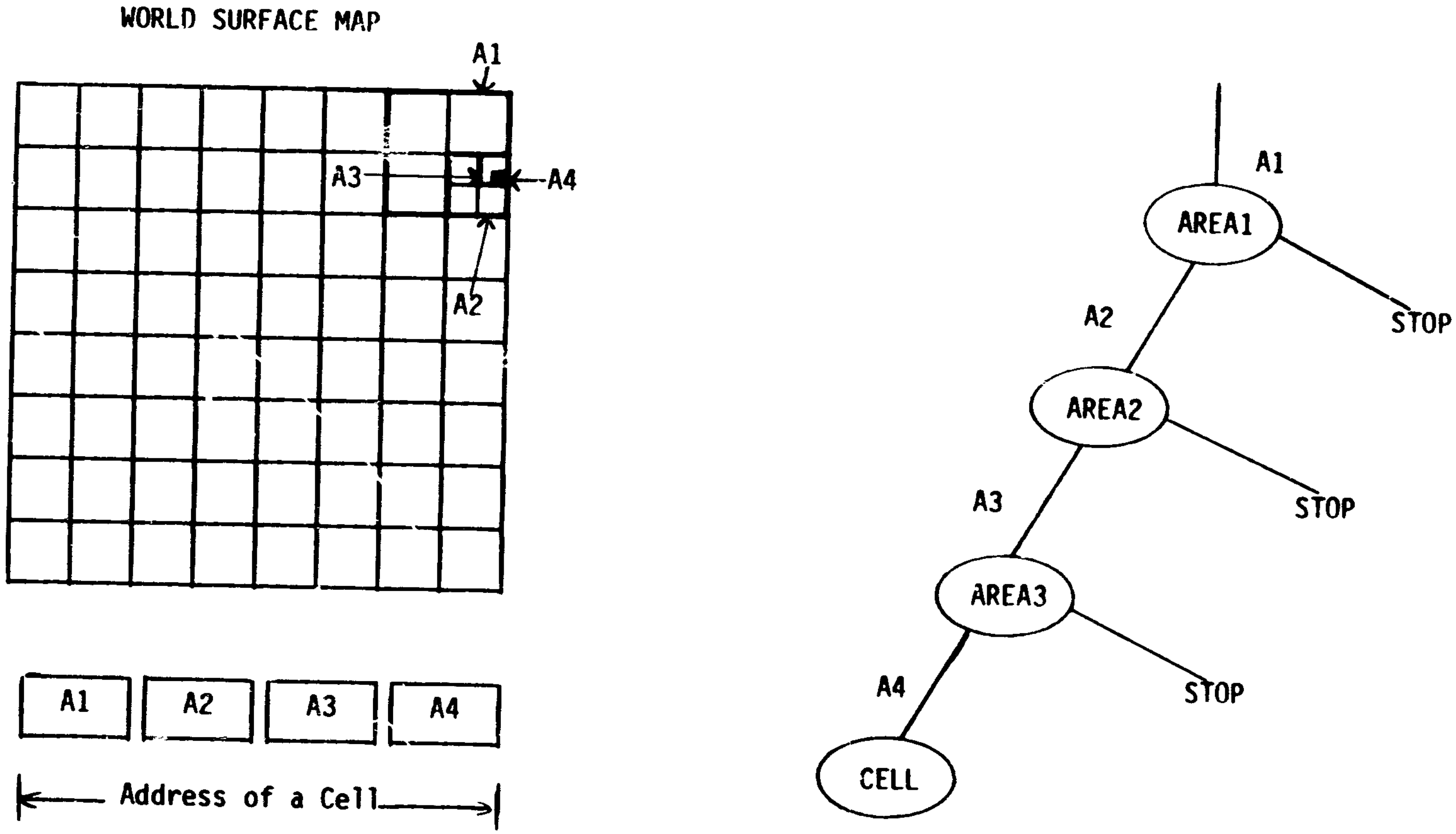


Figure 2-5

The Conceptual Data Retrieval Procedure by Telescoping

Four levels are used in this example to address a resolution cell of desired size. The tree-like structure representing the process is also shown.

2.2 Cost Analysis of Telescoping

There are two parameters to be determined for the telescoping retrieval procedure. The first parameter is the number of telescoping steps required to achieve 7 km resolution and the second is the area block size at each step. An optimization procedure is formulated to determine these parameters. It involves the formulation of cost functions at each and every telescoping level and the analysis of some empirical data.

Telescoping is based on the fact that there exists large homogeneous area blocks, that is, blocks consist of a single type of surface. It can be easily seen that the percentage of heterogeneous area blocks on the Earth's surface decreases with the area block size. Heterogeneous area blocks are area blocks that contain different surface types (for example, blocks contain land-ocean boundary). The relationship between the amount of heterogeneous area blocks on the Earth's surface and different block sizes is empirically determined. Figure 2-6 plots the data.

The x-axis, M, represents the number of partitions, 7 km apart, on a side of the area block and the y-axis, $f(M)$, represents the percentage of heterogeneous area blocks of the total Earth's surface area. The plot reveals a linear relationship between M and $f(M)$ when $M<400$. As M approaches zero, the curve intersects the y-axis at $f(M) = 0.1$. In other words, there are about 10% heterogeneous cells (i.e., land-ocean boundaries) on the Earth's surface when the resolution approaches 7 km. This linear relationship is incorporated into the optimization analysis described as follows.

Let n_j be the number of partitions on a side of a square area block at the j^{th} telescoping level, and N be the total number of 7 km partitions along the latitutde, or half the perimeter of the Earth. N can be written as:

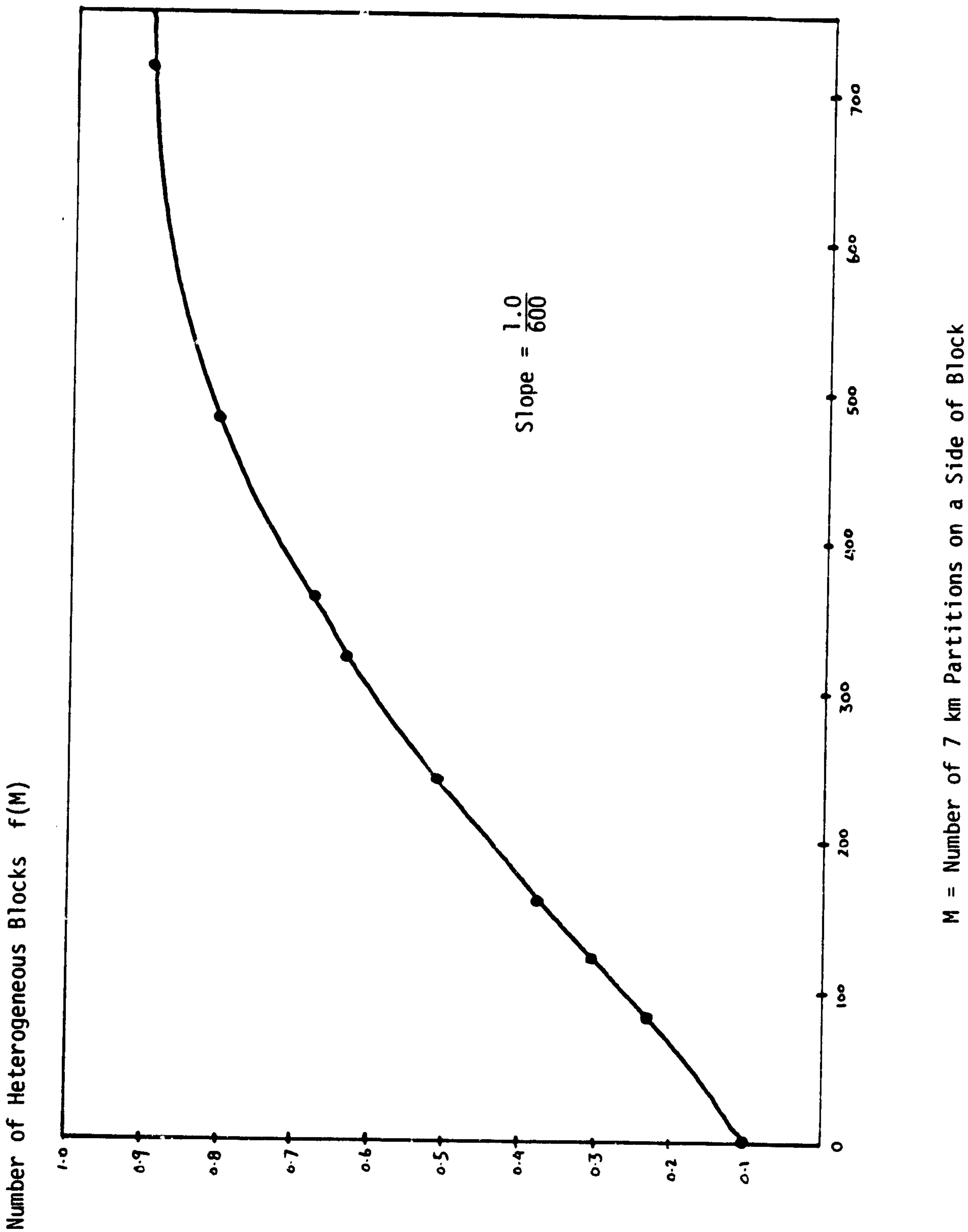


Figure 2-6 M vs. f(M)

$$N = n_1 n_2 \cdots n_i = \prod_{j=1}^{i} n_j$$

where i is the required number of telescoping levels to achieve 7 km resolution. N is a constant expressed as:

$$N = (\tfrac{1}{2})\ 2\pi R/7\text{km}$$

$$= 2859.75$$

where $R = 6371$ km, the Earth's radius.

Further define a function f describing the plot in Figure 2-6 by

$$f(x) = a + b\,x$$

where a is the intercept and b is the slope.

The fractional amount of heterogeneous area blocks encountered at the k^{th} telescoping levels can be expressed as:

$$f(\frac{N}{m_k}) = a + b(\frac{N}{m_k})$$

where $m_k = \prod_{j=1}^{k} n_j$.

The cost to telescope i levels is then expressed as:

$$C = \text{cost of level 1} + \text{cost of level 2} \cdots + \text{cost of level } i$$

$$C = n_1^2 + n_1^2 n_2^2 f(\frac{N}{m_1}) + \cdots + n_1^2 n_2^2 \cdots n_i^2 f(\frac{N}{m_1}) f(\frac{N}{m_2}) \cdots f(\frac{N}{m_{i-1}})$$

$$C = \sum_{r=1}^{i} [\prod_{j=1}^{r} n_j^2] [\prod_{k=i}^{r-1} f(\frac{N}{m_k})] .$$

To minimize the above cost function, the extrema have to be located. This involves the solving of the following set of equations for n_j .

$$\frac{\delta C}{\delta n_j} = 0 \qquad \text{where } j=1,\cdots i$$

To illustrate the process, a few simple cases of evaluating the n_j's are detailed below.

Case 1: $i=1$, implies no telescoping is applied.

$$N = n_1 \text{ , and } C = n_1^2$$

$$n_1 = 2860$$

$$C = 8.178 \times 10^6$$

Case 2: $i = 2$. Two levels of addressing

$$N = n_1 n_2 \text{ , and}$$

$$C = n_1^2 + n_1^2 n_1^2 f(\frac{N}{m_1})$$

$$= n_1^2 + n_1^2 n_1^2 (a+b\frac{N^3}{n_1})$$

To solve: $\frac{\delta C}{\delta n_1} = 0 = 2n_1 - b\frac{N^3}{n_1}$

The results: $n_1 = N(b/2)^{1/3} = 2C9$

$n_1 = (2/b)^{1/3} = 10.63$

$C = 1.103 \times 10^6$

<u>Case 3</u>: $i = 3$. Three levels of addressing

$N = n_1\ n_2\ n_3$, and

$$C = n_1^2 + n_1^2\ n_2^2\ f(\frac{N}{m_1}) + n_1^2\ n_2^2\ n_3^2\ f(\frac{N}{m_1})\ f(\frac{N}{m_2})$$

$$= n_1^2 + n_1^2\ n_1^2\ (a+b\frac{N}{n_1}) + n_1^2\ n_2^2\ n_3^2\ (a+b\frac{N}{n_1})\ (a+b\frac{N}{n_1 n_2})$$

To solve: $\frac{\delta C}{\delta n_1} = 0$

$\frac{\delta C}{\delta n_2} = 0$

$\frac{\delta C}{\delta n_3} = 0$

The results: $n_1 = [\frac{abn^3}{2} + \frac{3}{4}\frac{b^2N^3}{(b/2)^{1/3}}]^{1/3} = 135.11$

$n_2 = \frac{N}{n_1}\ (b/2)^{1/3} = 1.99$

$$n_3 = (2/b)^{1/3} \qquad = 10.63$$

$$C \qquad = 0.17\times10^6$$

Case 4: $i = 4$. Four levels of addressing

$$N = n_1\, n_2\, n_3\, n_4\, , \text{ and}$$

$$C = n_1^2 + n_1^2\, n_2^2\, f(\frac{N}{m_1}) + n_1^2\, n_2^2\, n_3^2\, f(\frac{N}{m_1})\, f(\frac{N}{m_2}) +$$

$$n_1^2\, n_2^2\, n_3^2\, n_4^2\, f(\frac{N}{m_1})\, f(\frac{N}{m_2})\, f(\frac{N}{m_3})$$

The results: $n_1 = 72.25$

$n_2 = 1.87$

$n_3 = 1.99$

$n_4 = 10.63$

$C = 0.034\times10^6$

To summarize, the following table lists results of the first four cases. Although the n_i's are optimally computed, they are re-assigned to integers which are powers of two. This enables the binary serial address to be used most efficiently. In all cases, the number of partitions at the last telescoping level is 12. The area block involved in the last level of partition can be considered as a logical record. In this case, the record size is 12 x 12 = 144.

i	Computed n_i				Reassigned n_i				Cost
1				2857.7				-	8.178×10^6
2			269.0	10.6			256	12	1.185×10^6
3		135.1	1.9	10.6		128	2	12	0.179×10^6
4	72.3	1.9	2.0	10.6	64	2	2	12	0.035×10^6

The cost of the five cases is illustrated graphically in Figure 2-7. Case 4 is recommended. It consists of four telescoping steps. At the first level, the World Surface Map is partitioned into 2 x 64 x 64 = 8192 square areas employing 13 bits of address. The second step partitions each and every one of the 8192 areas into four squares, adding another two bits to the address and resulting in a total of 4 x 8192 = 32,768 area blocks. The third telescoping level is the same as the second one. The last step addresses the 144 elements within a record and requires an address of eight bits. Figure 2-8 summarizes this recommended telescoping procedure.

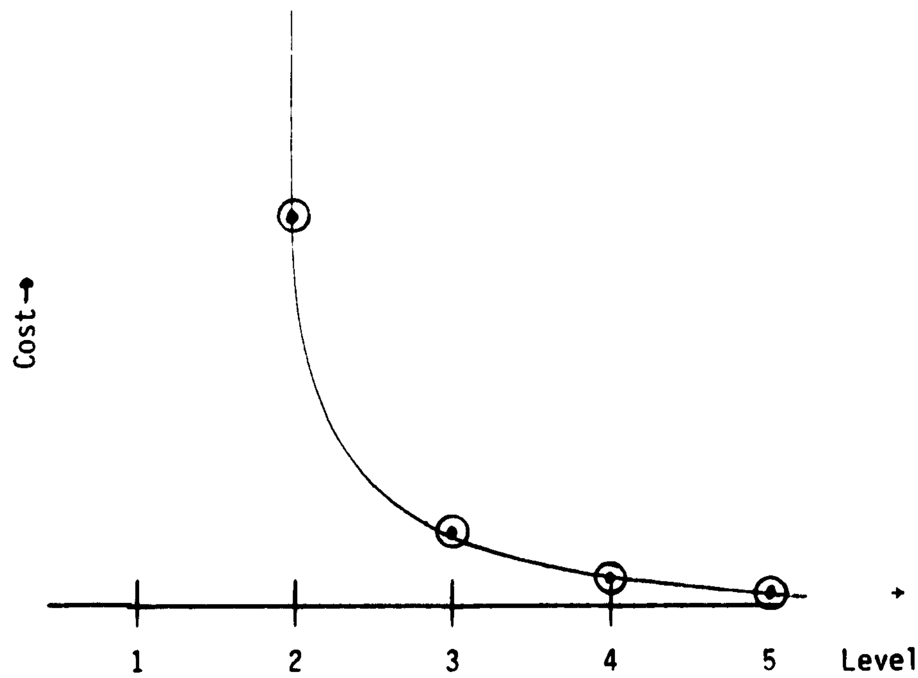


Figure 2-7 The Cost of Telescoping

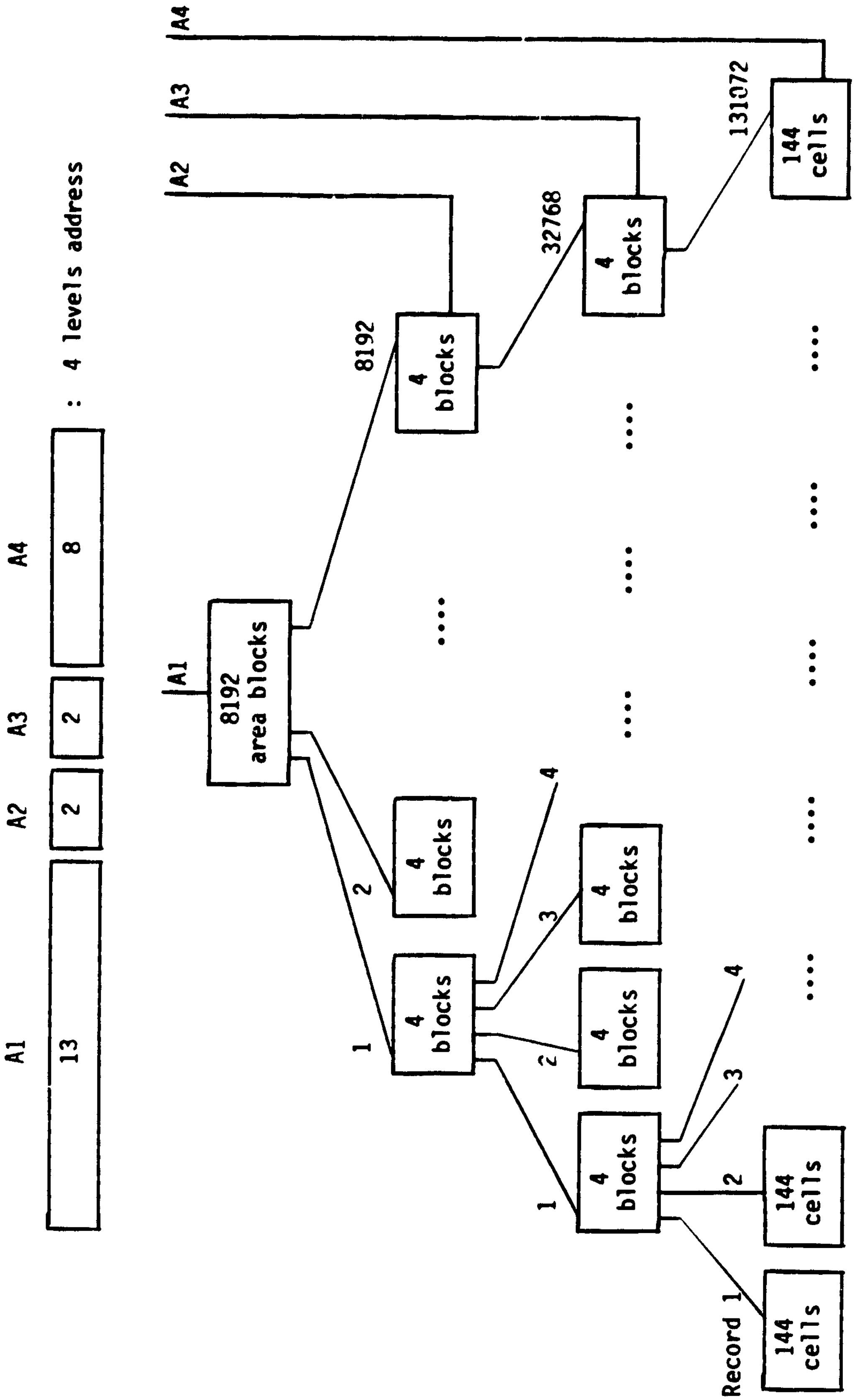


Figure 2-8 The Recommended Telescoping Structure

3.0 EXAMPLE OF A DATA BASE STRUCTURE BASED ON THE SPHERICAL ICOSAHEDRON

3.1 Introduction

The spherical icosahedron data base provides some useful properties which are valuable in many applications. In this model, the sphere is visualized as a spherical icosahedron as illustrated in Figure 3-1. This spherical icosahedron is obtained by radially projecting the edges of an inscribed icosahedron.

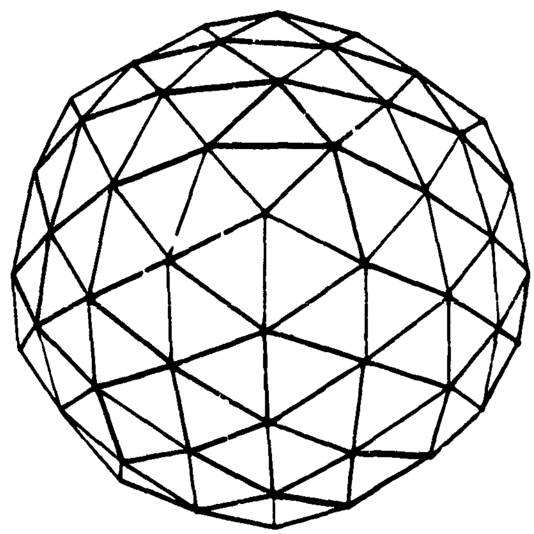
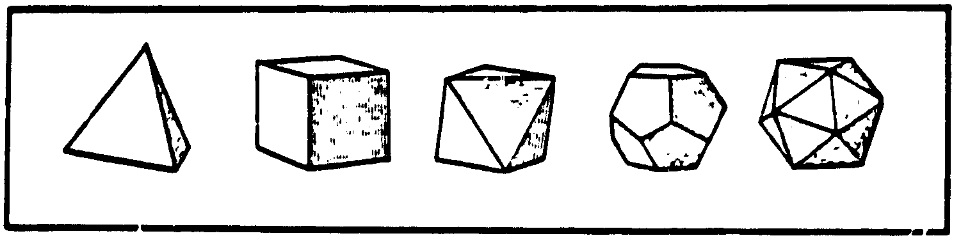


Figure 3-1 Spherical Icosahedron

3.2 Geometric Properties of a Spherical Icosahedron

A regular n-polyhedron is a geometric solid whose faces are regular n-sided polygon. It is well known that there can be only five regular polyhedrons as shown in Figure 3-2.



Regular Polyhedrons	Number and Type of Faces
Tetrahedron	4 equilateral triangles
Octahedron	8 equilateral triangles
Icosahedron	20 equilateral triangles
Hexahedron (cube)	6 squares
Dodecahedron	12 regular pentagons

Figure 3-2 The Five Regular Polyhedrons

Each regular polyhedron can be inscribed in a sphere, and the edges can be radially projected onto the sphere forming "spherical polyhedrons". The spherical icosahedron formed out of 20 spherical triangles is the division of the sphere having the greatest number of regular pieces. For the triangles of the spherical icosahedron, illustrated in Figure 3-3, the vertex angles have the value $\alpha = 2\pi/5$.

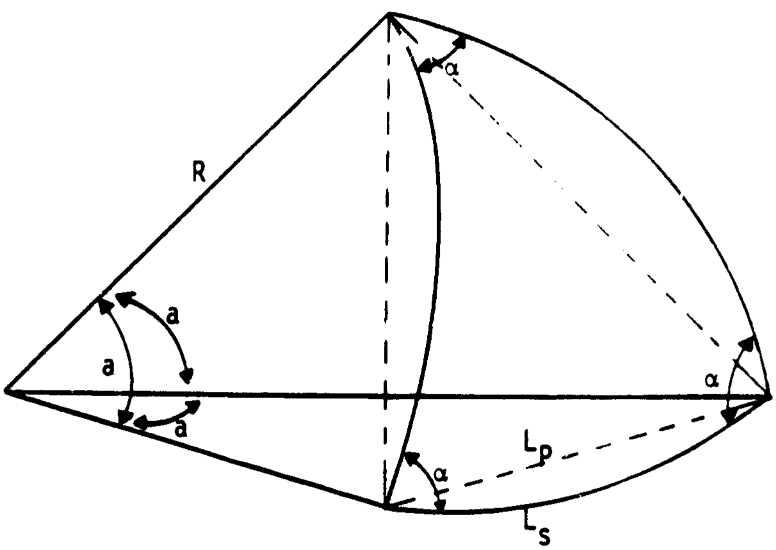


Figure 3-3 Face of Spherical Icosahedron

It can be shown that the side L_p of a plane equilateral triangle of the regular icosahedron is given by

$$L_p = 2R \sin \frac{a}{2}$$

where R is the radius of the sphere and "a" is the side of the spherical triangle given by:

$$a = 2 \cos^{-1} \left(\frac{\cos \frac{\alpha}{2}}{\sin \alpha}\right) \quad \text{or}$$

$$L_s = aR \ .$$

The area of the spherical triangle, A_s , and of the plane equilateral triangle, A_p , can be computed by:

$$A_s = \left(\frac{1}{20}\right) 4\pi R^2 \ , \text{ and}$$

$$A_p = \frac{\sqrt{3}}{4} L_p^2 \ .$$

Using 6371 kilometers as the radius of the Earth, the following list gives useful values involving the above equations.

$R = 6371$ km
$\alpha = 1.2566$
$a = 1.1071$
$L_p = 6698.8659$ km
$L_s = 7053.6527$ km
$A_p = 19431360.3$ km^2
$A_s = 25503223.6$ km^2

3.3 Spherical Icosahedron Data Base

In constructing a data base structured about the icosahedron solid, the process involves the division of the 20 spherical triangles into smaller triangles of nearly equal area. Consider the partitioning of the plane triangle of the regular icosahedron into equal-area equilateral triangles by successive processes of bisection of the sides. After four such bisection processes, the original triangles are partitioned into 64 smaller triangles as shown in Figure 3-4. Each additional bisection quadruples the number of partitions. The data base element can be defined in form of a hexagon and each hexagonal resolution element has twice the area of a small triangle.

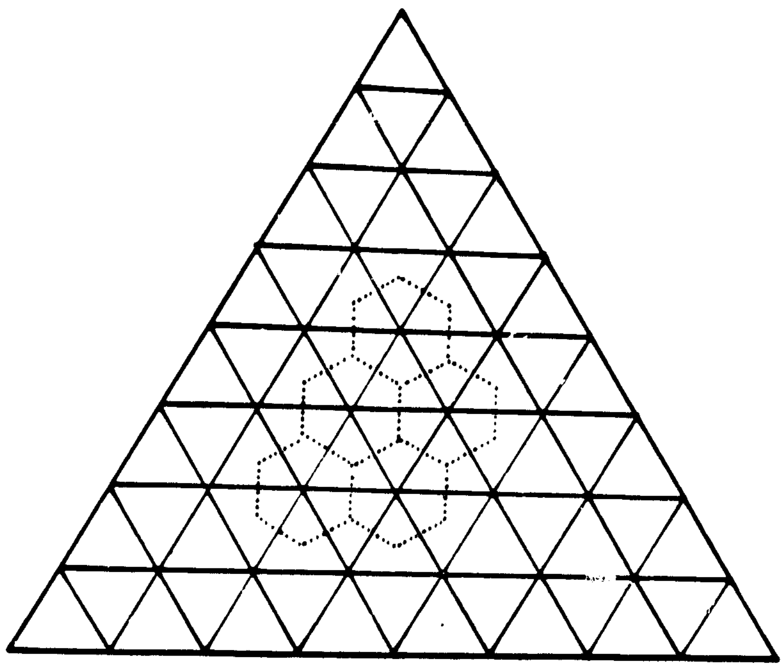


Figure 3-4 An Example of Partitioning an Icosahedron Triangle
(Resolution elements are defined as hexagons.)

Using the above example shown in Figure 3-4, a basic triangle of 7053.65 kilometers on all sides consists of 32 resolution elements at the resolution of about 881.7 kilometers in width and 796975.7 square kilometers in area. To approach the resolution of 7 kilometers, the fundamental length of the basic triangle has to be partitioned into 1024 parts. Each hexagonal resolution element corresponds to about 48.6 square kilometers. Figure 3-5 shows other alternatives possible for resolution element sizes.

From Figure 3-4, it is obvious that equal-area elements on the plane triangle do not radially project as equal-area elements on the spherical triangle. For example, those elements near the center of the plane triangle have larger projections than those elements near the edges and vertices of the plane triangle. Hence, if a triangular grid of equal-area elements is first constructed on the plane triangle, it is then necessary to distort this grid into a curvilinear network so that the elements near the center are smaller than those near the edges. The distortion is such that when the curvilinear elements are projected radially, equal-area elements are again obtained on the spherical surface. The desired sequence of transformation is illustrated in Figure 3-6. The mathematical details of deriving these transformations are discussed in the following section.

3.4 Mathematical Formulation of the Mapping Function

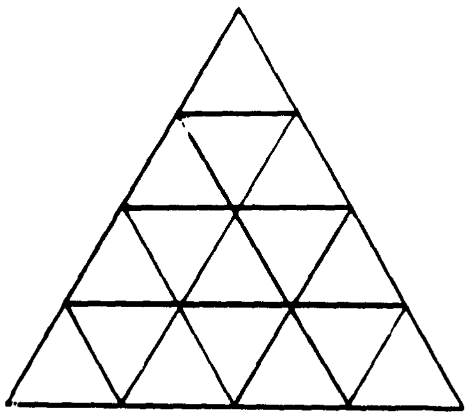
First, consider a plane surface subtended by a spherical surface with radius R. Let $\bar{r}_0$ be the vector from the center of the sphere to the given plane. As shown in Figure 3-7, let dA_p be an area element on this plane, and let $\bar{r}$ be the vector from the center of the sphere to the area element dA_p. Let dA_s be the spherical area element obtained by projected dA_p radially onto the sphere. Then, it is not too difficult to show that the area element dA_s on the sphere is related to the area element dA_p of the plane triangle via the relation:

SIDE PARTITIONS	NUMBER OF HEXAGONS (million)	WIDTH OF HEXAGONS	AREA OF THE HEXAGONAL ELEMENT
256	0.65	27.6 km	778.3 km^2
512	2.62	13.8 km	194.6
1024	10.48	6.9 km	48.6
2048	41.94	3.4 km	12.2
4096	167.77	1.7 km	3.0

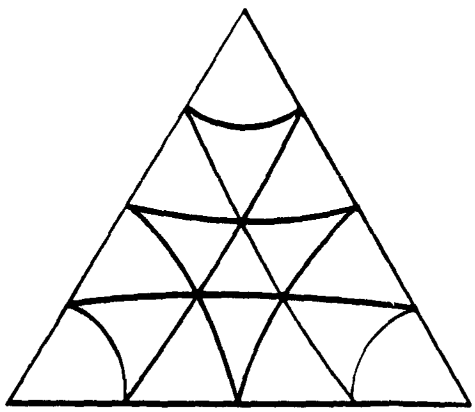
Notes:

- Side partition is the number of partitions of a side of a basic icosahedron triangle.
- Number of Hexagons is the total number of resolution cells on the Earth-oriented spherical surface and is computed by 1/2 (side partitions)^2 x 20.
- The width of a hexagon is measured as the shortest distance between two parallel sides.
- The area of a hexagonal element is computed by (Total Earth's Surface Area) (Number of Hexagons)

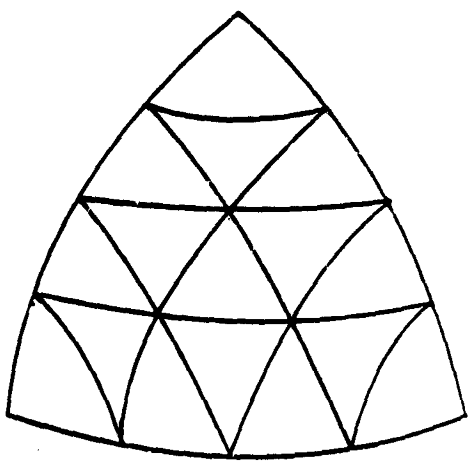
Figure 3-5 Various Resolution Sizes of the Icosahedron Data Structure



Plain Triangles with Equal Areas by Partitioning



Plain Triangles with Distorted Grids by Mapping



Spherical Triangles with Equal Areas by Projecting

Figure 3-6 Sequence of Transformation

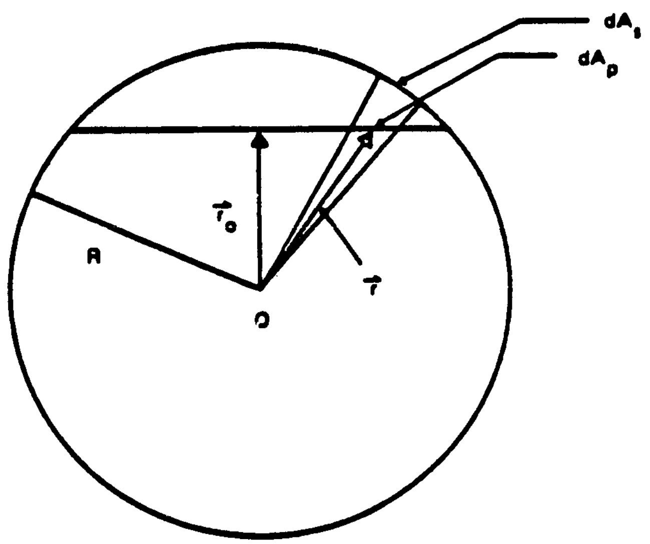


Figure 3-7 Relation Between Plane and Spherical Area Elements

$$dA_s = \frac{R^2 \cos^3 (\bar{r},\bar{r}_o)}{r_o^{\,2}} dA_p$$

where $(\bar{r},\bar{r}_o)$ denotes the angle between $\bar{r}$ and $\bar{r}_o$.

This relation is used to derive the mapping function in transforming to equal-area elements on the sphere. The mapping function is expressed in a new coordinate system (u,v,w) defined as lines of constant interval forming the partitioning of the larger triangle (Figure 3-8). In particular, $u=0$, $v=0$, $w=0$ are lines forming the boundary of the triangle; $u=1$, $v=1$, $w=1$ correspond to the three vertices. Let $\bar{r}$ be a simple vector from the center of the icosahedron to any point on the triangle and it can be expressed as:

$$\bar{r} = u\bar{A} + v\bar{B} + w\bar{C}$$

where $\bar{A}$, $\bar{B}$, and $\bar{C}$ are vectors from the center to the three vertices of the triangle containing $\bar{r}$. $\bar{r}$ can also be expressed in terms of the mapping function $g(u,\rho)$

$$\bar{r} = \frac{g(u,\rho)\bar{A} + g(v,\rho)\bar{B} + g(w,\rho)\bar{C}}{g(u,\rho) + g(v,\rho) + g(w,\rho)}$$

where

$$g(u,\rho) = \frac{1}{3} + \gamma(u-\frac{1}{3}) -3(1-\gamma)\ (u-\frac{1}{3})^2$$

$$- \frac{3}{2}\ (1-\gamma)\rho + \frac{3}{23}\ u(u-\frac{1}{3})^2$$

$$\rho = (u-\frac{1}{3})^2 + (v-\frac{1}{3})^2 + (w-\frac{1}{3})^2$$

and

$$\gamma^2 \simeq 0.8290$$

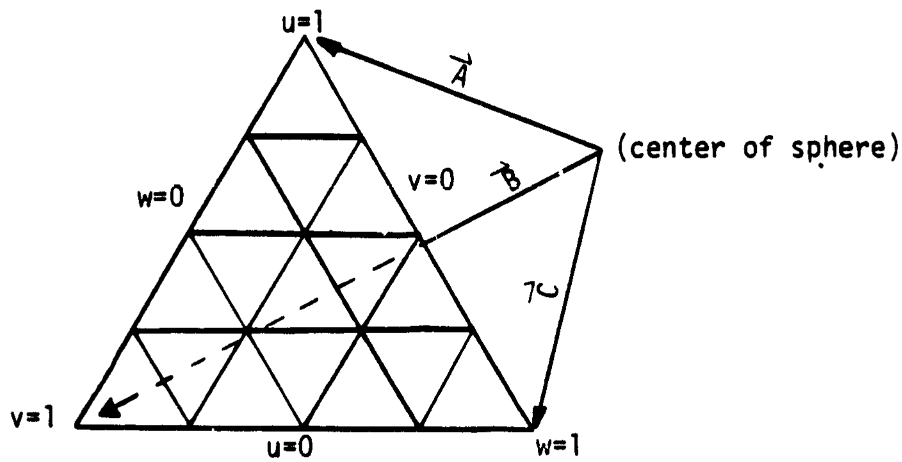


Figure 3-8 Coordinate System in a Plane Triangle

Using this transformation, a data base can be constructed by projecting, onto the sphere, the vertices of triangles obtained by successive bisection of the faces of the icosahedron. These points are centers of nearly equal-area hexagon-like resolution elements. The 12 vertices of the icosahedron are centers of pentagons.

3.5 Organization of Data Base and Addressing

The data base system adopted from the near equal-area division of the spherical icosahedron is most likely to be organized into a binary division scheme. The major objective is to provide "telescoping", that is, each finer level of resolution is contained in the previous resolution. The addressing scheme being binarily indexable is capable of accessing each finer level of resolution.

The mapping of the data base onto storage array may be accomplished in a number of ways. The classic method is mapping by columns and rows, as is done in the usual mathematical matrix notation. This column-row mapping may cause problems in the spherical icosahedron data base because of the nature of the data base structure - triangular structure. Other mapping techniques include scatter storage with one-to-one mapping and various consistent logical stringing techniques.

The specific scheme studied in this report is based on binary division scheme. This mapping starts at the level of the faces of the spherical icosahedron, numbering these faces 1 through 20 as in Figure 3-9. Each face is divided, to finer resolution levels, by a triangular grid, as shown in Figure 3-10. On each level of division, the areas are divided into quadrants, which are labeled by a 2-bit binary number. Each level of division, N , is indicated by the addition of two binary bits to the least-significant end of a 5+N bit binary number. This binary address is the serial location of a triangular area in the 20×2^{2N} elements location array. An address to the third level of division would appear in Figure 3-11. Figure 3-12 provides the possible levels of divi-

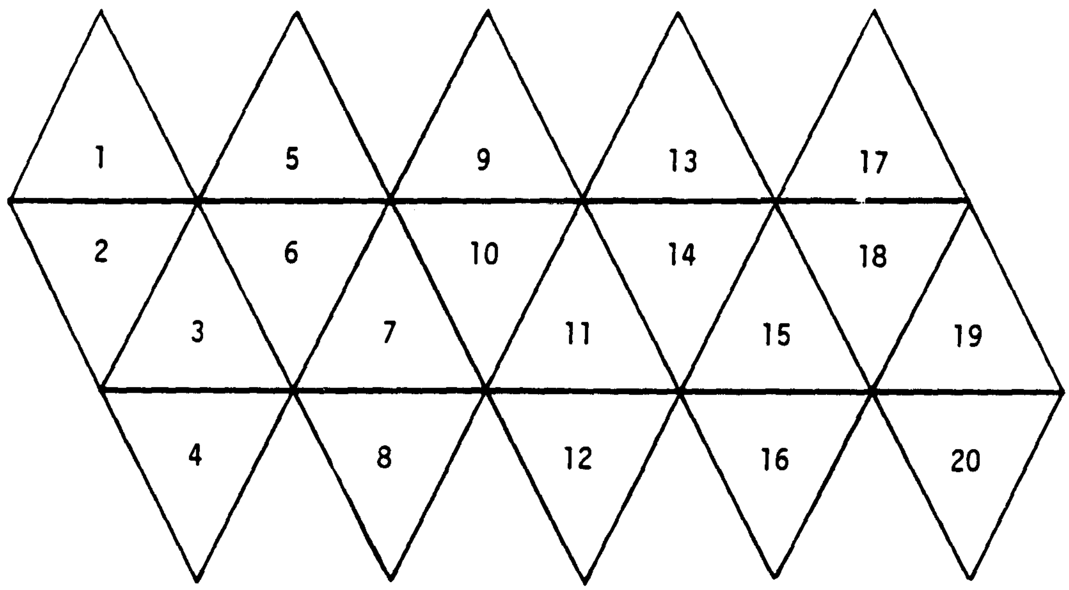


Figure 3-9 Icosahedron Face Numbering Scheme
(A 5-bit string is required, from 00000 to 10011)

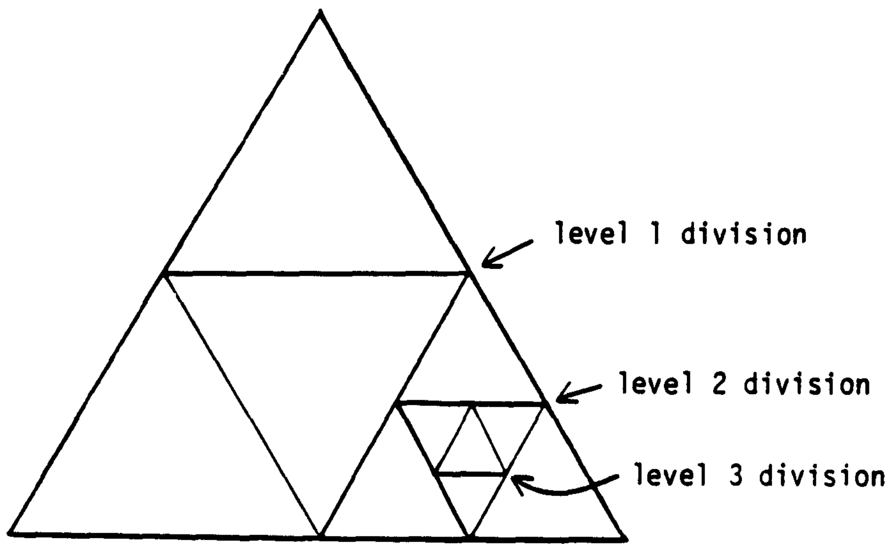


Figure 3-10 Binary Division Scheme

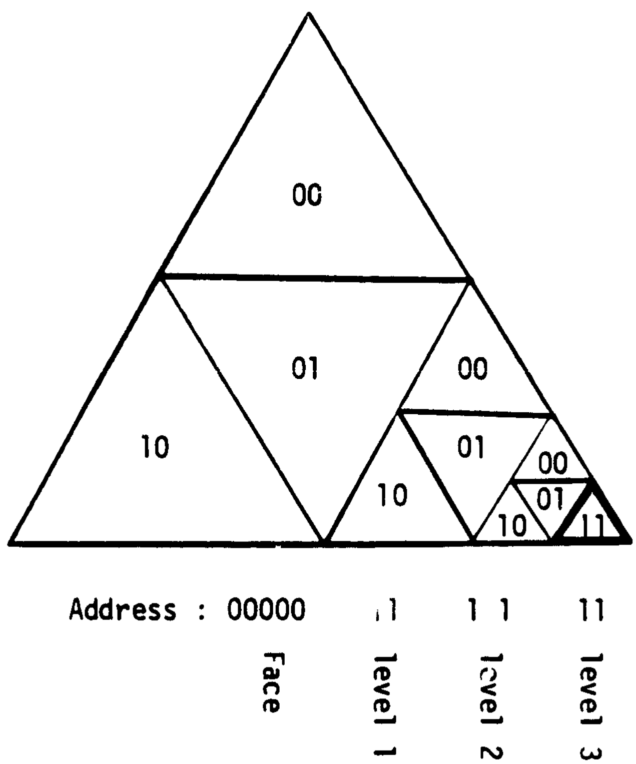


Figure 3-11 Example of Addressing by Binary Bits

LEVEL OF DIVISIONS	SIDE PARTITIONS	LENGTH OF THE SMALLEST TRIANGLE	NUMBER OF ADDRESSING BITS REQUIRED
1	2	3527.0	5 + 2
2	4	1763.5	5 + 4
3	8	881.8	5 + 6
4	16	440.9	5 + 8
5	32	220.4	5 + 10
6	64	110.2	5 + 12
7	128	55.1	5 + 14
8	256	27.6	5 + 16
9	512	13.8	5 + 18
10	1024	6.9	5 + 20
11	2048	3.4	5 + 22
12	4096	1.7	5 + 24

Note: The five bits are for addressing the 20 icosahedron triangles, from 00000 to 10011. The rest of the address (N bits) is for addressing the 2^{2N} triangular elements in the basic icosahedron triangle.

Figure 3-12 Level of Division vs. Addressing Requirement sions, their corresponding sizes (lengths of the triangles) and their corresponding address requirements.

In developing the method of ordering data in the data base, there are obvious advantages of using a serial addressing scheme rather than a normal two-dimensional addressing scheme. The expression of addresses as a single bit string allows storage of addresses as single machine words, whereas a two-dimensional addressing scheme would require two or three words, including one for the face number. Also, the expandibility and generality of the serial string permits the use at any resolution level without the recalculation of the serial index.

The serial addressing string can be constructed simply by considering the high order m bits as the record number, and the lower (N-M) bits as the address within record. The record number points to the triangular area containing the desired resolution element and address of the hexagonal cell is represented by the lower bit string. A specific example of addressing a 7 km resolution cell is provided below. Figure 3-13 illustrates the addressing procedure of a hexagonal resolution element and its corresponding address. The first five bits of the serial string is the face number indexing one of the 20 icosahedron triangles. The next 12 bits represent one of the 4096 smaller triangles after six levels of division. The small triangle has a side of 110.2 km and is considered as a record. The last seven bits is the address of one of the 128 resolution cells within the record in a predefined serial fashion. One way to number the resolution cells is illustrated in Figure 3-13. The hexagonal resolution cell has a width of approximately 7 km. Two of the corner resolution cells and two of the elements are stored in an adjacent record. As shown in Figure 3-13, only half of the remaining side resolution elements are associated with and stored in the logical record. Thus, with this scheme, there is no duplication of storage. Other possibilities exist for the indexing of hexagonal resolution elements forming logical records. There are $2(4^n-1)$possible alternatives are where n is the number of elements.

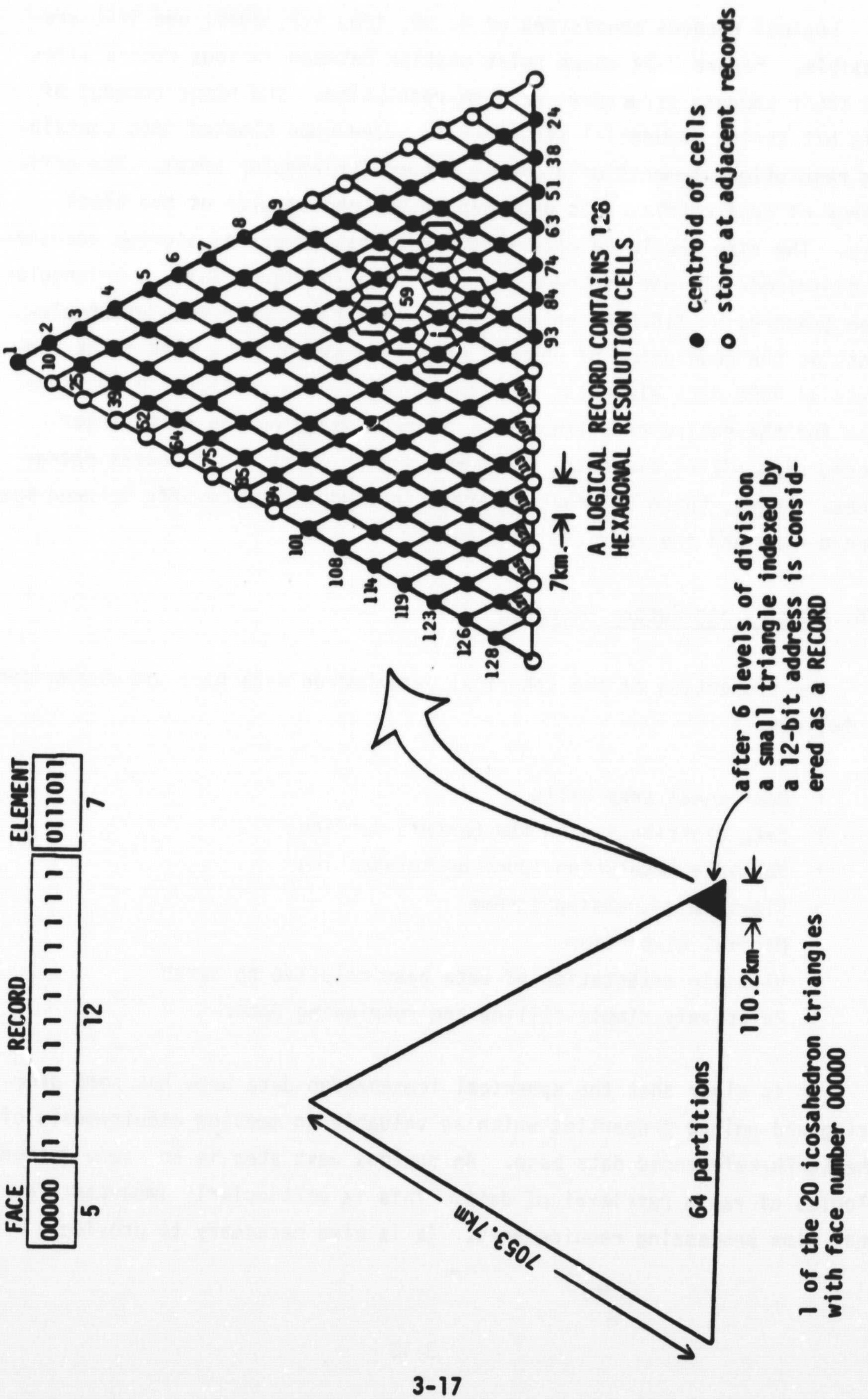


Figure 3-13 Addressing a Hexagonal Resolution Element

Logical records consisting of 8, 32, 128, 512, 2048, and 8192 are possible. Figure 3-14 shows relationships between various record sizes and their address structures at 7 km resolution. The basic concept of this bit string sequential structure is based upon blocked data containing resolution elements of a single record (triangular area). The efficiency of data retrieval is depended on the chosen size of the block data. The size should be determined such that number of storage accesses is minimized. In one of the cases as listed in Figure 3-14, a triangular area (record) of 110.2 km on each side consists of 8192 resolution elements at the resolution of about 7 km. If a sector of a disk track consists of 4096 data elements, only two accesses are required to retrieve data for the entire area (record). Storage based on the next larger record size, 32768 elements, would require at least eight access operations. Hence, there is great interest in studying trade-offs between the record size and the retrieval efficiency.

3.6 Summary and Future Research

The properties of the spherical icosahedron data base are summarized as follows:

- Near-equal area cells
- Even distribution on the Earth's surface
- Variable resolution level by telescoping
- Flexible addressing scheme
- Minimal distortion
- Flexible orientation of data base relative to Earth
- Relatively simple filling and retrieving data.

It is clear that the spherical icosahedron data base has some promising and unique properties which ae valuable in meeting requirements of the Earth-referenced data base. An obvious next step is to study methodologies of rapid retrieval of data. This is particularly important for real time processing requirements. It is also necessary to provide a

FACE NUMBER	RECORD NUMBER	ADDRESS OF RESOLUTION ELEMENTS	RECORD SIZE
5	16	3	8
5	14	5	32
5	12	7	128
5	10	9	512
5	8	11	2048
5	6	13	8192
5	4	15	32768

Figure 3-14 Relationships Between Record Sizes and Address Structures at 7 Kilometers Resolution

means of rapid transformation between the geographic coordinates and data base structural coordinaes for addressing purposes. Also, demonstrations will be in order to validate all the proposed procedures including the retrieval process, the filling method, and the serially transferring process of data to a "rectangular" configurated storage memory.

3.7 References

[1] Beaudet, P. R., F. K. Chan and L. Goldshalk, "Organizational Structures for Constant Resolution Earth Data Bases", Computer Sciences Corporation, 6002-1 (prepared for the Environmental Prediction Research Facility, Monterey, CA), November 1973.

[2] Chan, F. K. and E. M. O'Neill, "Feasibility Study of a Quadrilateralized Spherical Cube Earth Data Base", CSC TR-75-6007 (prepared for the Environmental Prediction Research Facility, Monterey, CA), March 1975.

[3] O'Neill, E. M. and R. E. Laubscher, "Extended Studies of a Quadrilateralized Spherical Cube Earth Data Base", CSC NEPRF, TR 3-76 (prepared for the Environmental Prediction Research Facility, Monterey, CA), May 1976.

[4] Maling, D. H., Coordinate Systems and Map Projections, George Philip and Son Limited, London, 1973.